Systems and Technology Group

The Power Within the Cell Processor - and How to Unleash It

David J. Krolak
Senior Engineer
IBM Systems and Technology Group
Rochester, Minnesota

Alex Chow
Manager, IBM Cell Design Center
IBM Systems and Technology Group
Austin, Texas

Systems and Technology Group

The Element Interconnect Bus of the Broadband Engine Processor

David J. Krolak
Senior Engineer
IBM Systems and Technology Group
Rochester, Minnesota

Agenda

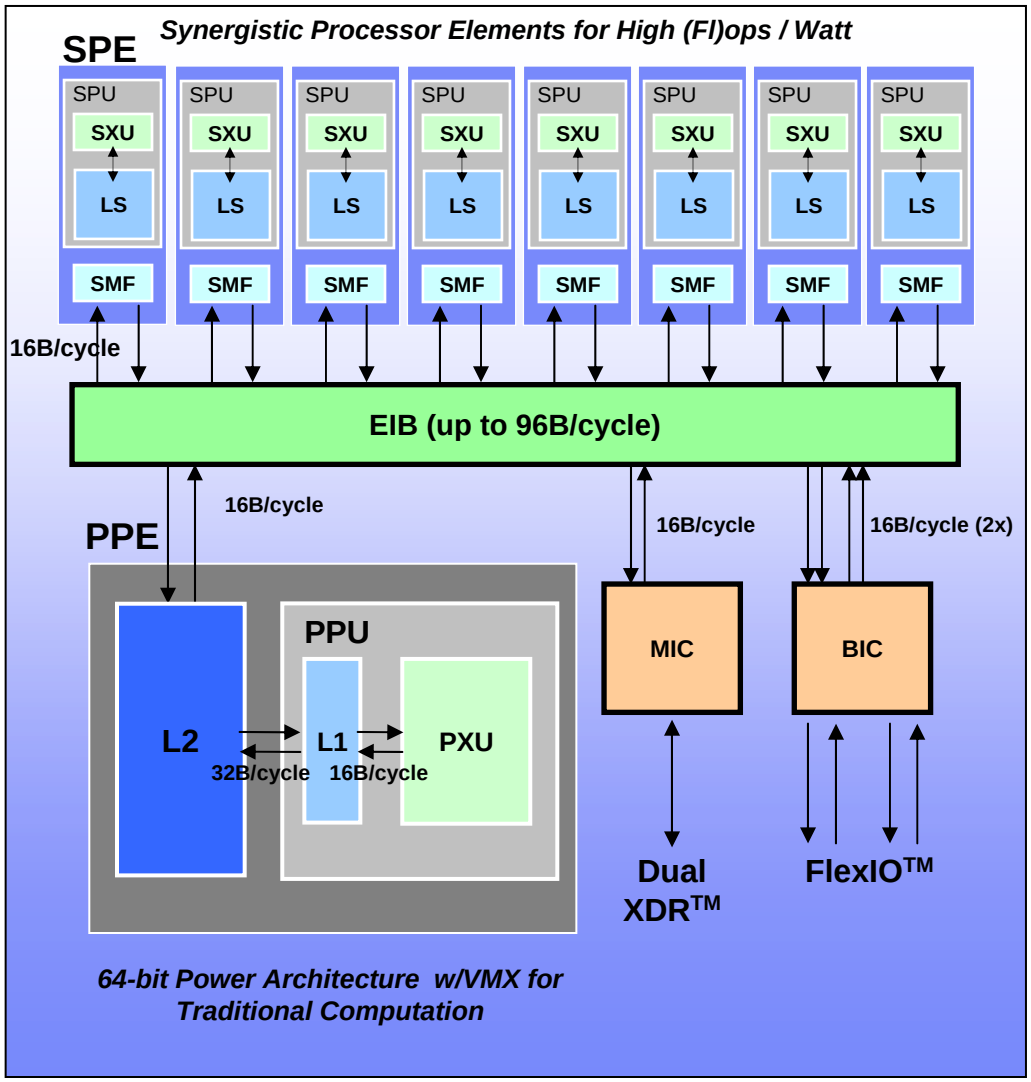
- **Cell Overview from an EIB perspective**
- **Key Attributes of the EIB**
- **Software Considerations**
- **Summary**

Cell is a demanding environment for an on-chip interconnect to support

- **Supercomputer on a chip**
- **Multi-core microprocessor (9 cores)**
- **3 to 4 GHz core clock frequency**
- **10x performance for many applications**
- **Digital home to distributed computing**

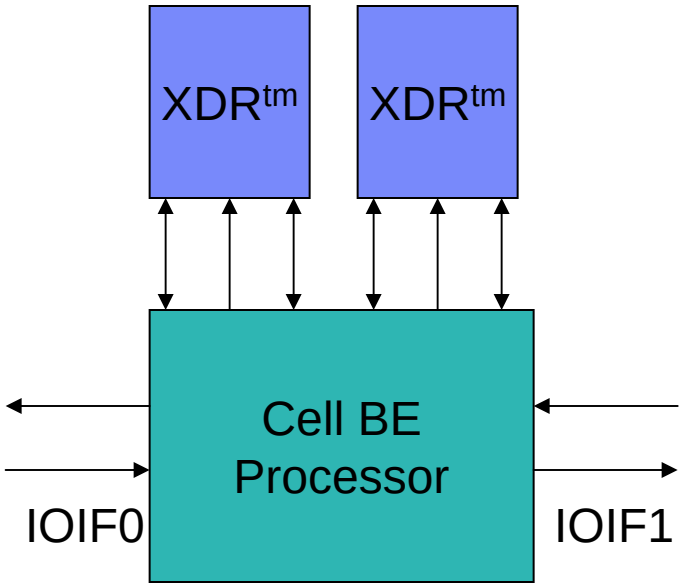
Cell Broadband Engine Key Features

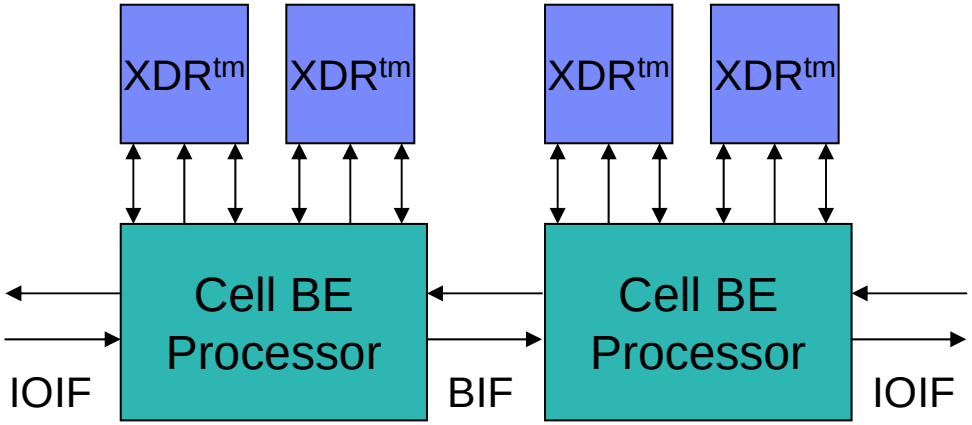
- **The first generation CELL processor consists of:**
  - A Power Processor Element (PPE)
  - 8 Synergistic Processor Elements (SPE) with Synergistic Memory Flow Control (SMF)
  - A high bandwidth Element Interconnect Bus (EIB)
  - A Bus Interface Controller with two configurable I/O interfaces (BIC)
  - A Memory Interface Controller (MIC)

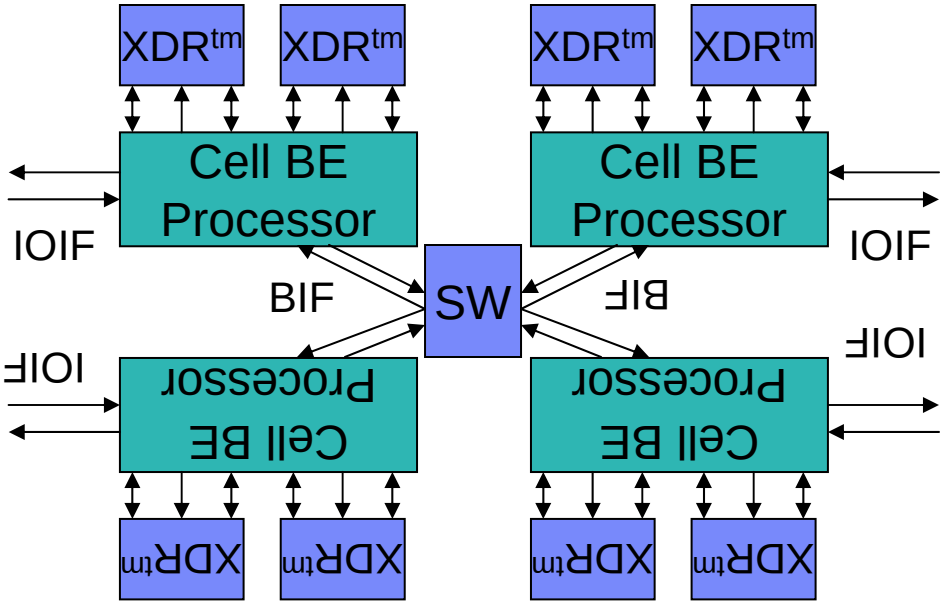


Cell BE Processor Can Support Many Systems

- Game console systems
- Blades
- HDTV
- Home media servers
- Supercomputers



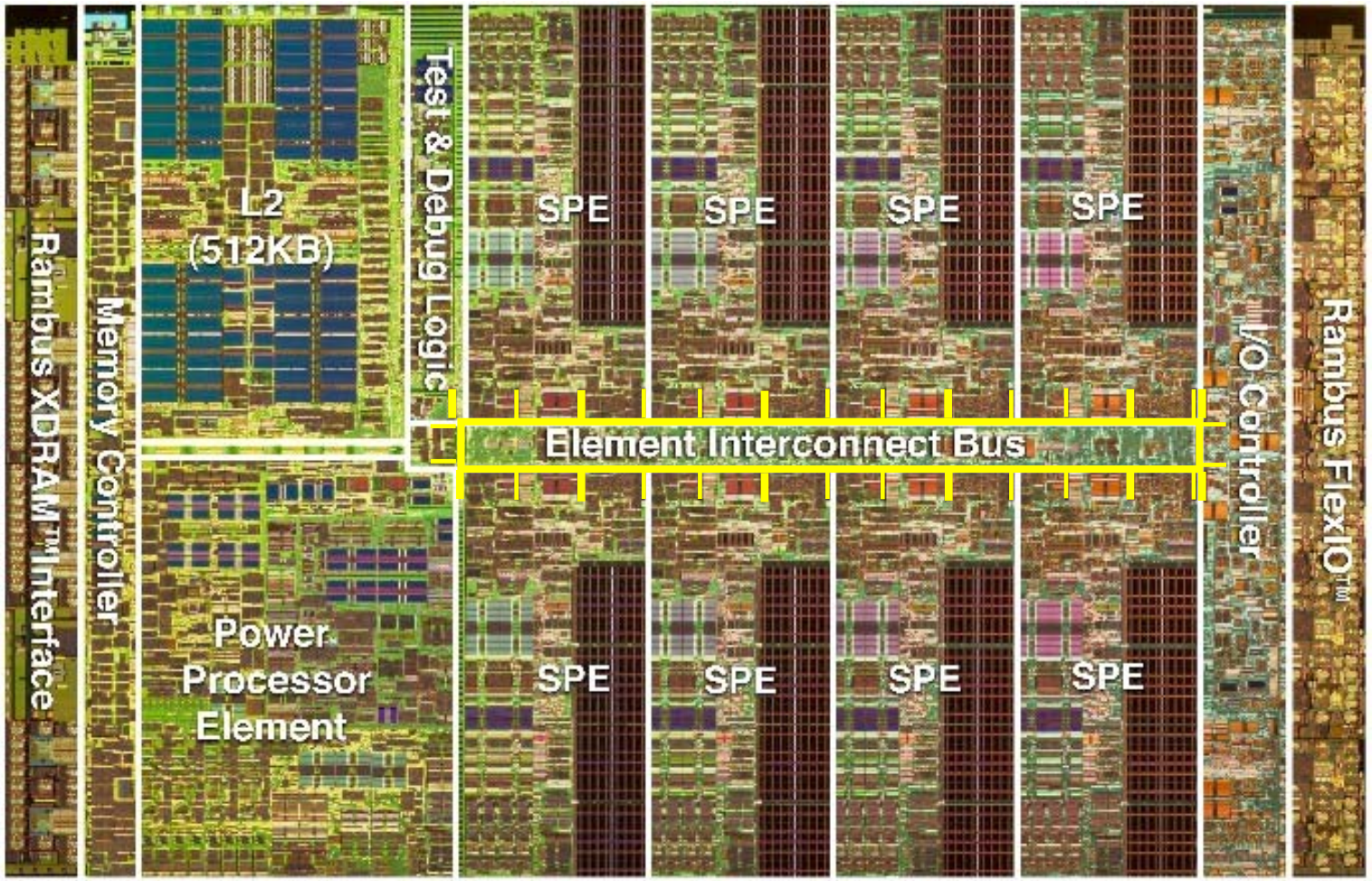




Element Interconnect Bus Overview

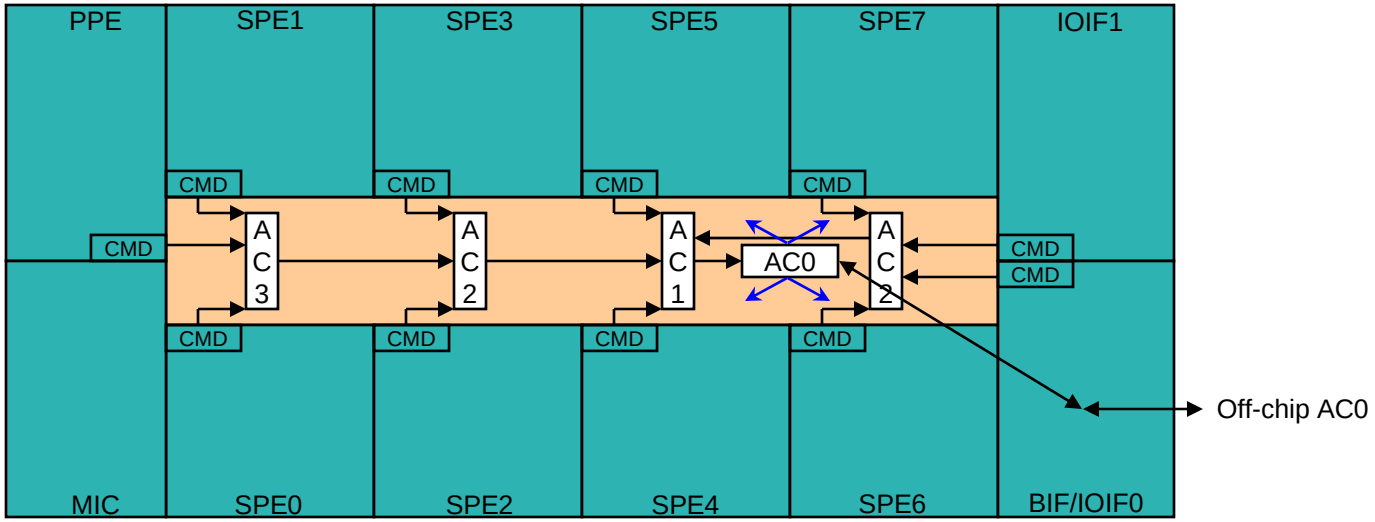
- **Coherent SMP Bus**
  - Supports 64 outstanding requests per requestor
  - Address collision detection and prevention
- **Independent Command and Data Networks**
- **Split Command / Data Transactions**
- **High Bandwidth**
  - Four 16 Byte data rings
  - Operates at ½ processor core frequency
  - Up to 96 Bytes / processor cycle → 192 Bytes / bus cycle
    - Peak rate 300 GB/s at 3.2 GHz processor clock; up to 200GB/s sustained
    - 16 Bytes / bus cycle source and 16 Bytes / bus cycle sink per port
    - 12 Element ports
- **Modular Design for Scalability**
  - Physical modularity for flexibility

Element Interconnect Bus



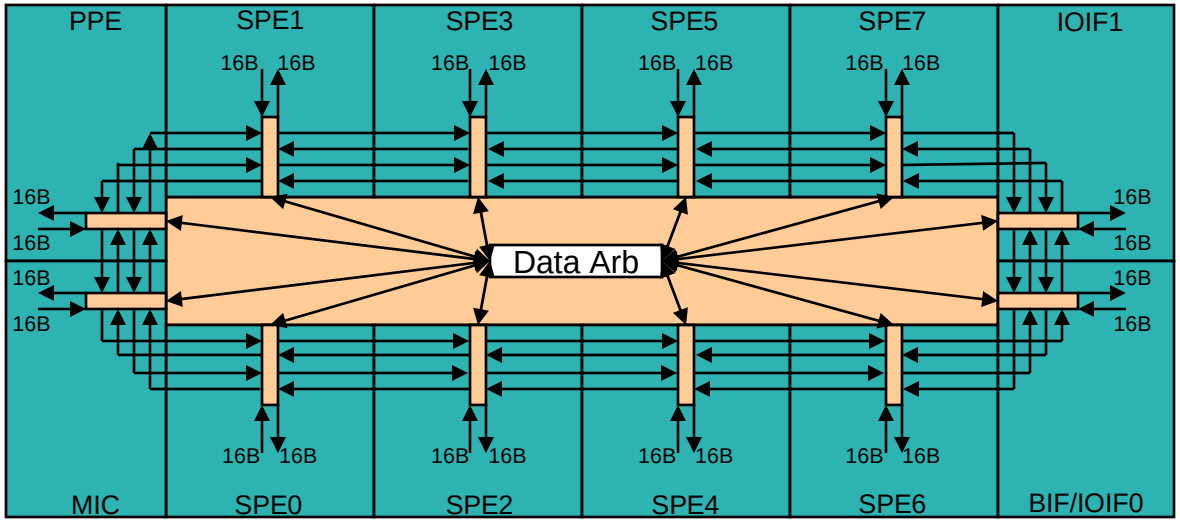
Element Interconnect Bus – Command Topology

- **“Address Concentrator” tree structure minimizes wiring resources**
- **Single serial command reflection point (AC0)**
- **Address collision detection and prevention**
- **Fully pipelined**
- **Content –aware round robin arbitration**
- **Credit-based flow control**



Element Interconnect Bus - Data Topology

- **Four 16B data rings connecting 12 bus elements**
  - Two clockwise / Two counter-clockwise
- **Physically overlaps all processor elements**
- **Central arbiter supports up to three concurrent transfers per data ring**
  - Two stage, dual round robin arbiter
- **Each element port simultaneously supports 16B in and 16B out data path**
  - Ring topology is transparent to element data interface



Next Generation I/O Bandwidth

- **Today, a typical "Front Side Bus" has available data bandwidth of 6 to 8 Gigabytes/sec**
- **Typical DDR2 Memory interfaces run in the 6 to 11 GB/sec range**

In contrast, the Cell has:

- **Two Rambus I/O controllers**
  - **BIF/IOIF0**
    - **Configurable protocol**
      - **Broadband Engine Interface (BIF) coherent protocol**
      - **I/O Interface (IOIF) non-coherent protocol**
    - **Scalable from 0 to 6 bytes outbound / 0 to 5 bytes inbound**
      - **Up to 30 GB/s outbound / 25 GB/s inbound in 5 GB/s increments**
  - **IOIF1**
    - **IOIF protocol**
    - **Scalable from 0 to 2 bytes outbound / 0 to 2 bytes inbound**
      - **Up to 10 GB/s outbound / 10 GB/s inbound in 5 GB/s increments**
- **A Rambus Dual XDR™ Memory Controller which supports an aggregate memory data bandwidth of 25.6 GB/s.**

Superlative internal bandwidth capability

- **Each EIB Bus data port supports 25.6GBytes/sec* *in each direction***
- **The EIB Command Bus streams commands fast enough to support 102.4 GB/sec for coherent commands, and 204.8 GB/sec for non-coherent commands.**
- **The EIB data rings can sustain 204.8GB/sec for certain workloads, with transient rates as high as 307.2GB/sec between bus units**

Despite all that available bandwidth…

* The above numbers assume a 3.2GHz core frequency – internal bandwidth scales with core frequency

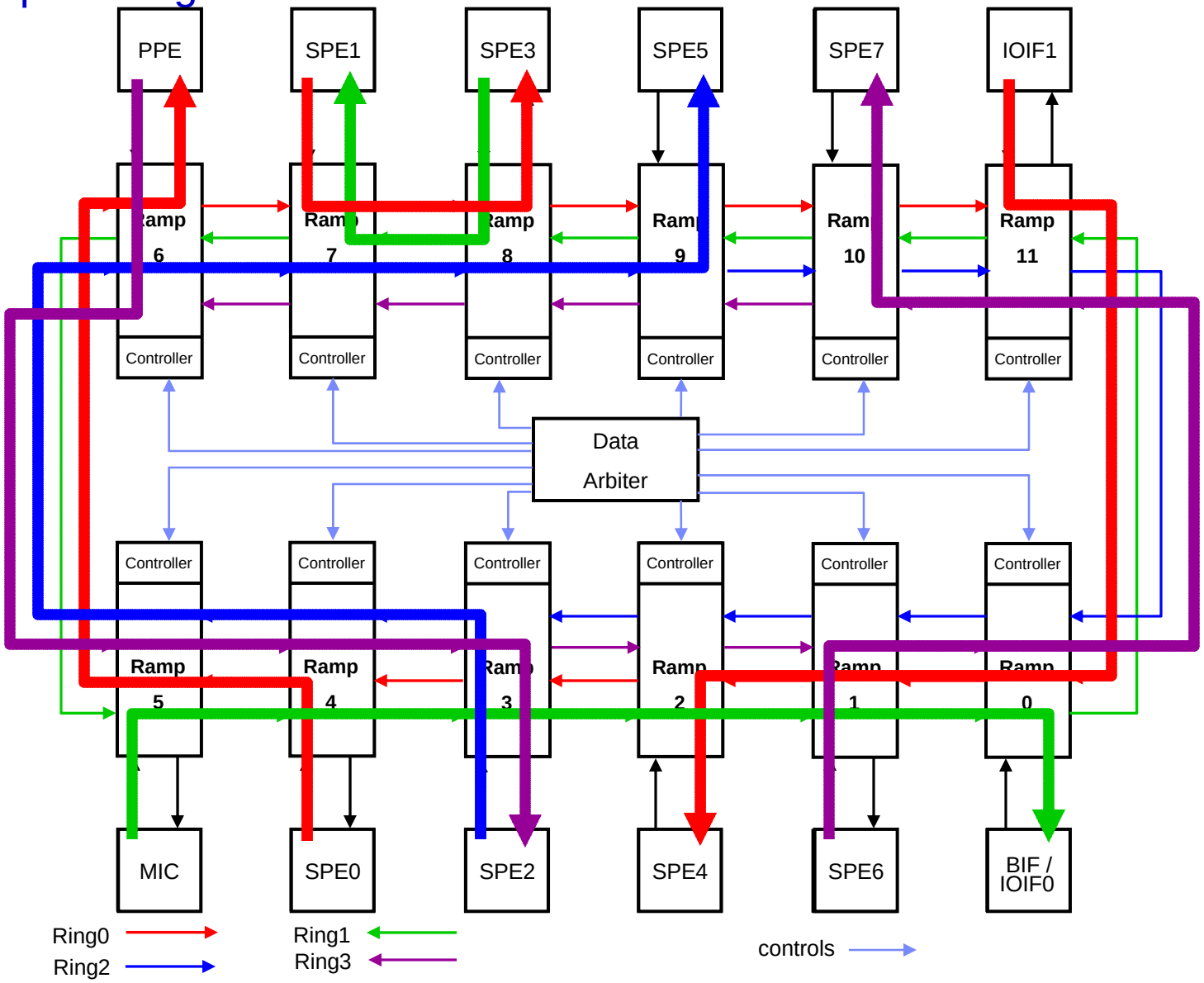
There are some minor bottlenecks

- **The data rings are a shared resource**
- **Multiple transactions can coexist on one ring, but a transfer between two units will block access to that ring for any other units along the path**

….So it can pay to manage which devices talk to each other when configuring workload assignments.

That said, it usually handles ring conflicts pretty well:

Example of eight concurrent transactions



…And some “bigger” bottlenecks –

Like the bandwidth of an off-chip port vs. the on-chip potential demand for it:

- **For example, if only four SPEs tried to move a block of data to or from the XDR memory at the same time, the demand would be 102GB/s*, vs. the capability of the XDR interface to provide at best only 25.6 GB/s.**
- **The XDR memory unit would quickly be overwhelmed, and start to retry commands.**
  …Potentially causing unrelated traffic to slow down as well.
- **Similar imbalances can exist with contention for an individual memory bank, and for the two I/O ports**

This situation can be controlled with…

*SPE demand assumes a 3.2GHz core frequency – internal bandwidth scales with core frequency

Resource Allocation Management

- **Optional facility used to minimize over-allocation effects of critical resources**
  - Independent but complementary function to the EIB
  - Critical (managed) resource's time is distributed among groups of requestors
- **Managed resources include:**
  - Rambus XDR™ DRAM memory banks (0 to 15)
  - BIF/IOIF0 Inbound and BIF/IOIF0 Outbound
  - IOIF1 Inbound and IOIF1 Outbound
- **Requestors Allocated to Four Resource Allocation Groups (RAG)**
  - 17 requestors – PPE, SPEs, I/O Inbound (4 VCs), I/O Outbound (4 VCs)
- **Central Token Manager controller**
  - Requestors ask permission to issue EIB commands to managed resources
  - Tokens granted across RAGs allow requestor access to issue command to the EIB
  - Round robin allocation within RAG
  - Dynamic software configuration of the Token Manager to adjust token allocation rates for varying workloads
  - Multi-level hardware feedback from managed resource congestion to throttle token allocation

Summary

- **The Cell chip with its Element Interconnect Bus ushers in unprecedented data bandwidth capability for digital media and entertainment**
- **However, workload assigments and access to critical elements such as memory and I/O must be managed by software in order to obtain maximum performance**
- **Cell is projected to enable entirely new classes of applications, even beyond what we can contemplate today**

IBM Microelectronics Division
1580 Route 52, Bldg. 504
Hopewell Junction, NY 12533-6351

The IBM home page is http://www.ibm.com
The IBM Microelectronics Division home page is
http://www.chips.ibm.com